RAW SEQUENCE LISTING ERROR REPORT
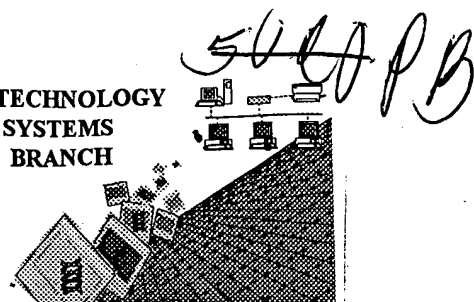
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/049,626_

Source: _PCT/10_

Date Processed by STIC: _3/7/2002_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
> http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   > Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

ERROR DETECTED   **SUGGESTED CORRECTION**   SERIAL NUMBER: _10/049,626_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino Numbering

The numbering under each 5ᵗʰ amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length

Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) ____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 ____ Skipped Sequences (OLD RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences (NEW RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.

10 ____ Invalid <213> Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial-Sequence

11 ____ Use of <220>

Sequence(s) ____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

PCT10

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/049,626

DATE: 03/07/2002
TIME: 15:37:21

Input Set : A:\pto.vsk.txt
Output Set: N:\CRF3\03072002\J049626.raw

pp 1-3

delete

Does Not Comply
Corrected Diskette Needed

W--> 1 (? ? ?)
5 <110> APPLICANT: Director-General of Agency of Industrial Science and Technology
6   Nishimatu Construction CO.,LTD
7   NYK LOGISTICS TECHNOLOGY INSTITUTE
9 <120> TITLE OF INVENTION: A method for detecting or quantifying bacteria having a
specific function
10   and the genes thereof from natural environment, novel 16 S rDNA gene
11   information, and probes
13 <130> FILE REFERENCE: probe
15 <140> CURRENT APPLICATION NUMBER: US/10/049,626
16 <141> CURRENT FILING DATE: 2002-02-22
18 <150> PRIOR APPLICATION NUMBER: JP P1999-237818
19 <151> PRIOR FILING DATE: 1999-08-25
21 <160> NUMBER OF SEQ ID NOS: 9
23 <170> SOFTWARE: PatentIn Ver. 2.0

ERRORED SEQUENCES

62 <210> SEQ ID NO: 2   *1528 shown (p. 2)*
63 <211> LENGTH: (1538)
64 <212> TYPE: DNA
65 <213> ORGANISM: Cycloclasticus pugetii
67 <220> FEATURE:
68 <221> NAME/KEY: rRNA
69 <222> LOCATION: (1)..(1538)
71 <400> SEQUENCE: 2
72 agagtttgat catggctcag attgaacgct ggcggcatgc ctaacacatg caagtcgaac 60
73 ggaaacgatg ctagcttgct agcaggcgtc gagtggcgga cgggtgagta atgcatagga 120
74 atctacctaa tagtgtggga caacctggtg aaaaccaggc taataccgca taatccctac 180
75 ggggcaaagc aggggacctt cgggccttgc gctaatagat gagcctatgt cggattagct 240
76 agttggtgag gtaatggctc accaaggcaa cgatccgtag ctggtttgag aggatgatca 300
77 gccacactgg gactgagaca cggcccagac tcctacggga ggcagcagtg gggaatattg 360
78 cacaatggag gaaactctga tgcagcaatg ccgcgtgtgt gaagaaggcc ttaggggttgt 420
79 aaagcacttt cagtagggag gaaaagttta aggttaataa ccttaggccc tgacgttacc 480
80 tacagaagaa gcaccggcta actccgtgcc acagccggcg gtaatacgga gggtgcaagc 540
81 gttaatcgga attactgggc gtaaagcgcg cgtaggcggt taaacaagtc agatgtgaaa 600
82 gccccgggct caacctggga actgcatttg aaactgttta gctagagtgt ggtagaggag 660
83 agtggaattt caggtgtagc ggtgaaatgc gtagatatct gaaggaacac cagtggcgaa 720
84 ggcggctctc tggaccaaca ctgacgctga ggtgcgaaag cgtgggtagc aaacgggatt 780
85 agataccccg gtagtccacg ccgtaaacga tgtcaactga ctgttgggcg ggtttccgct 840
86 tagtggtgca staacgcaat aagttgaccg cctggggagt acggccgcaa ggctaaaact 900
87 caaatgaatt gacggggggcc cgcacaagcg gtggagcatg tggtttaatt cgatgcaacg 960
88 cgaagaacct tacctaccct tgacatacag agaactttct agagatagat tggtgccttc 1020

3/7/02

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/049,626

DATE: 03/07/2002
TIME: 15:37:21

Input Set : A:\pto.vsk.txt
Output Set: N:\CRF3\03072002\J049626.raw

```
 89 gggaactctg atacaggtgc tgcatggctg tcgtcagctc gtgtcgtgag atgttgggtt 1080
 90 aagtcccgta acgagcgcaa cccttatcct tagttgctac catttagttg ggcactctaa 1140
 91 ggagactgcc ggtgataaac cggaggaagg tggggacgac gtcaagtcat catggccctt 1200
 92 atgggtaggg ctacacacgt gctacaatgg ccggtacaga gggccgcaaa ctcgcgagag 1260
 93 taagctaatc ccttaaagcc ggtcctagtc cggattgcag tctgcaactc gactgcatga 1320
 94 agctggaatc gctagtaatc gcggatcaga atgccgcggt gaattcgttc ccgggccttg 1380
 95 tacacaccgc ccgtcacacc atgggagtgg gttgcaaaag aagtgggtag ctaacttcg 1440
 96 ggaggccgct caccactttg tgattcatga ctggggtgaa gtcgtaacaa ggtagcccta 1500
```

E--> 97 gggggaacctg gggctggatc acctcctt   (1538) 1528

```
 99 <210> SEQ ID NO: 3
100 <211> LENGTH: (1532)  1529
101 <212> TYPE: DNA
102 <213> ORGANISM: Cycloclasticus pugetii
104 <220> FEATURE:
105 <221> NAME/KEY: rRNA
106 <222> LOCATION: (1)..(1532)
108 <400> SEQUENCE: 3
109 agagtttgat catggctcag attgaacgct ggcggcatgc ctaacacatg caagtcgaac 60
110 ggaaacgatg ctagcttgct agcaggcgtc gagtggcgga cgggtgagta atgcatagga 120
111 atctacctaa cagtggggga caacctggtg aaaaccagsc taataccgca taatccctaa 180
112 cgggcaaagc aggggacctt cgggccttgc gctaatagat gagcctatgt cggattagct 240
113 agttggtgag gtaatggccc accaaggcaa cgatccgtag ctggtttgag aggatgatca 300
114 gccacactgg gactgagaca cggcccagac tcctacggga ggcagcagtg gggaatattg 360
115 cacaatggag gaaactctga tgcagcaatg ccgcgtgtgt gaagaaggcc ttagggttgt 420
116 aaagcacttt cagtagggag gaaaagttta aggttaataa ccttaggccc tgacgttacc 480
117 tacagaagaa gcaccggcta actccgtgcc agcagccgcg gtaatacgga ggggtgcaag 540
118 cgttaatcgg aattactggg cgtaaagcgc gcgtaggcgg ttaaacaagt cagatgtgaa 600
119 agccccgggc tcaacctggg aactgcattt gaaactgttt agctagagtg tggtagagga 660
120 gagtggaatt tcaggtgtag cggtgaaatg cgtagatatc tgaaggaaca ccagtggcga 720
121 aggcggctct ctggaccaac actgacgctg aggtgcgaaa gcgtgggtag caaacgggat 780
122 tagatacccc ggtagtccac gccgtaaacg atgtcaacta actgttgggc gggtttccgc 840
123 ttagtggtgc astaacgcaa taagttgacc gcctggggag tacggccgca aggctaaaac 900
124 tcaaatgaat tgacggggggc cgcacaagc ggtggagcat gtggtttaat tcgatgcaac 960
125 gcgaagaacc ttacctaccc ttgacataca gagaactttc tagagataga ttggtgcctt 1020
126 cgggaactct gatacaggtg ctgcatggct gtcgtcagct cgtgtcgtga gatgttgggt 1080
```

W--> 127 taagtcccgt aacgagcgcⒶ Ⓝcttatcct tagttgctac catttagttg ggcactctaa 1140
```
128 ggagactgcc ggtgataaac cggaggaagg tggggacgac gtcaagtcat catggccctt 1200
129 atgggtaggg ctacacacgt gctacaatgg ccggtacaga gggccgcaaa ctcgcgagag 1260
130 taagctaatc ccttaaagcc ggtcctagtc cggattgcag tctgcaactc gactgcatga 1320
131 agctggaatc gctagtaatc gcggatcaga atgccgcggt gaattcgttc ccgggccttg 1380
132 tacacaccgs ccgtcacacc atgggagtgg gttgcaaaag aagtgggtag ctaaccttc 1440
133 gggaggccgc tcaccacttt gtgattcatg actggggtga agtcgtaaca aggtagccct 1500
```

E--> 134 aggggaacct ggggctggat cacctcctt   (1539) 1529

```
217 <210> SEQ ID NO: 9
218 <211> LENGTH: 18
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial Sequence
222 <220> FEATURE:
223 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
```

p.3

see item 9 on Error Summary Sheet

RAW SEQUENCE LISTING   DATE: 03/07/2002
PATENT APPLICATION: US/10/049,626   TIME: 15:37:21

Input Set : A:\pto.vsk.txt
Output Set: N:\CRF3\03072002\J049626.raw

    225 <400> SEQUENCE: 9
    226 aaaggaggtg atccagcc   18
E--> 227 (1/1) delete

        Use of n and/or Xaa has been detected in the Sequence Listing.
        Review the Sequence Listing to insure a corresponding
        explanation is presented in the <220> to <223> fields of
        each sequence using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/049,626

DATE: 03/07/2002
TIME: 15:37:22

Input Set : **A:\pto.vsk.txt**
Output Set: **N:\CRF3\03072002\J049626.raw**

```
L:1 M:259 W: Allowed number of lines exceeded, (1) GENERAL INFORMATION:
L:15 M:270 C: Current Application Number differs, Replaced Application Number
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:97 M:254 E: No. of Bases conflict, LENGTH:Input:1538 Counted:1528 SEQ:2
L:97 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1538 Found:1528 SEQ:2
L:127 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3
L:127 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:1080
L:134 M:254 E: No. of Bases conflict, LENGTH:Input:1539 Counted:1529 SEQ:3
L:134 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1532 Found:1529 SEQ:3
L:160 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:4
L:160 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:840
L:227 M:254 E: No. of Bases conflict, LENGTH:Input:1 Counted:19 SEQ:9
L:227 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:2
L:227 M:252 E: No. of Seq. differs, <211> LENGTH:Input:18 Found:19 SEQ:9
```